A technique using principal component analysis to compare seasonal cycles of Earth radiation from CERES and model computations

G. Louis Smith,[1,2] Pamela E. Mlynczak,[2] and Gerald L. Potter[3]

Received 16 December 2011; revised 5 March 2012; accepted 28 March 2012; published 11 May 2012.

[1] A method for quantitatively comparing the seasonal cycles of two global data sets is presented. The seasonal cycles of absorbed solar radiation (ASR) and outgoing longwave radiation (OLR) have been computed from an eight-year data set from the Clouds and Earth's Radiant Energy System (CERES) scanning radiometers and from a model data set produced by the NASA Goddard Space Flight Center Global Modeling and Assimilation Office. To compare the seasonal cycles from these two data sets, principal component (PC) analysis is used, where the PCs express the time variations and the corresponding empirical orthogonal functions (EOFs) describe the geographic variations. Ocean has a long thermal response time compared to land, so land and ocean are separated for the analysis. The root-mean square values for the seasonal cycles of ASR and OLR are extremely close for the two data sets. The first three PCs are quite close, showing that the time responses and magnitudes over the globe are very similar. The agreement between the two sets of PCs is quantified by computing the matrix of inner products of the two sets. For ASR over land, the first PCs of CERES and the model agree to better than 99.9%. The EOF maps are similar for most of the globe, but differ in a few places, and the agreement of the EOF maps is likewise quantified. Maps of differences between the annual cycles show regions of agreement and disagreement.

Citation: Smith, G. L., P. E. Mlynczak, and G. L. Potter (2012), A technique using principal component analysis to compare seasonal cycles of Earth radiation from CERES and model computations, *J. Geophys. Res.*, *117*, D09116, doi:10.1029/2011JD017343.

1. Introduction

[2] Earth radiation budget observations are important because the climate machine is a heat engine, for which absorbed solar radiation (ASR) is the heat source and outgoing longwave radiation (OLR) is the heat sink. One application of Earth radiation budget measurements is comparison with model results. Numerical models incorporate our understanding of the processes that determine our weather and climate. Comparison of model results with measurements validates the processes of the model and may identify limits of the descriptions. The model must describe the transport of energy and moisture correctly in order to compute the correct distribution of OLR. Also, the model must accurately describe the clouds to determine the correct distribution of both ASR and OLR. The processes take place on a wide range of time and space scales, so it is necessary to make comparisons of model results at a variety of time scales, including the seasonal cycles, comprised of the annual and semiannual cycles, etc. Validations of both the diurnal cycles and seasonal cycles of a model are needed, as the processes of importance to these cycles differ considerably. Diurnal processes are local and mostly take place in the lower layer of the atmosphere and at the surface except for deep convection. After several days the effect of a change at any point is transmitted over the Earth, so that seasonal processes are globally connected.

[3] Many studies have been made to compare Earth radiation budget measurements with model results. *Bony et al.* [1992] used harmonic analysis to compare the seasonal cycles of cloud radiative forcing in the Laboratoire de Météorologie Dynamique circulation model with Earth Radiation Budget Experiment (ERBE) results. They also examined variations of zonal means of radiation fluxes in the time domain. *Kiehl et al.* [1998] compared the Earth's radiation budget as simulated by the NCAR Community Climate Model 3 with ERBE data in the time domain for selected regions. *Yang et al.* [1999] presented time variations of zonal means of TOA radiation to compare ERBE and NCEP/NCAR model monthly mean results. Many other comparisons have been performed between models and

[1]NASA Langley Research Center, Hampton, Virginia, USA.
[2]Science Systems and Applications, Inc., Hampton, Virginia, USA.
[3]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

Corresponding author: G. L. Smith, NASA Langley Research Center, Mail Stop 420, Hampton, VA 23681-2199, USA. (george.l.smith@nasa.gov)

0148-0227/12/2011JD017343

Table 1. Global Averages of Annual Mean Fluxes (W m^{-2})

	EBAF	Model	Model – EBAF	RMS of Difference
ASR	240.3	241.4	1.1	13.7
OLR	239.8	242.9	3.1	8.1

ERBE data for cloud radiative forcing. *Harrison et al.* [1990] and *Cullather et al.* [1997] used four months for the seasonal description of cloud radiative forcing. Others used summer and winter months for the comparison [*Barker et al.*, 1994; *Chen and Roeckner*, 1996; *Lin and Zhang*, 2004; *Cess et al.*, 1997]. *Potter and Cess* [2004] compared cloud radiative forcing results of ERBE with 19 models for regions during DJF. *Taylor et al.* [2011] used CERES data to show variations of zonal means through the seasonal cycle and to compare with the NCAR Community Climate System Model version 3 results in a study of radiation feedbacks.

[4] This paper presents a method for quantitatively comparing the seasonal cycles of two global data sets in the time domain by use of principal component analysis. To demonstrate the technique, a data set based on satellite observations from the CERES (Clouds and the Earth's Radiant Energy System) instruments [*Barkstrom*, 1990; *Wielicki et al.*, 1996; *Smith et al.*, 2011] is compared with model results from the Goddard Earth Observing System (GEOS) Atmospheric General Circulation Model Version 5 (GEOS-5). Monthly mean maps of ASR and OLR from the model and from CERES for the period March 2000 through August 2007 are used. These two data sets (based on satellite observations and model results) include parameters ASR and OLR that vary in time and space. The purpose of the model is to simulate numerically the dynamic processes that govern the ASR and OLR, so that the time response is a fundamental aspect of the model results and of the CERES observations. The quantitative comparison of the two data sets must contrast the time response observed with that computed by the model. The most efficient description of the time variation is obtained by use of principal component analysis (PCA). For a general discussion of PCA, please see *Wilks* [1995].

[5] *Mlynczak et al.* [2011] examined the seasonal cycles of Earth radiation measured by CERES with PCA and found that the annual cycle of ASR accounts for more than 95% of the overall variance of the seasonal cycles, which include the annual cycle, the semi-annual cycle, out-of-phase annual cycle, and higher-frequency terms. The use of principal components permits the reduction of the twelve monthly mean maps to a single map and is a major simplification for attempting to understand the time and space variability of the radiation fields. The method presented here can be

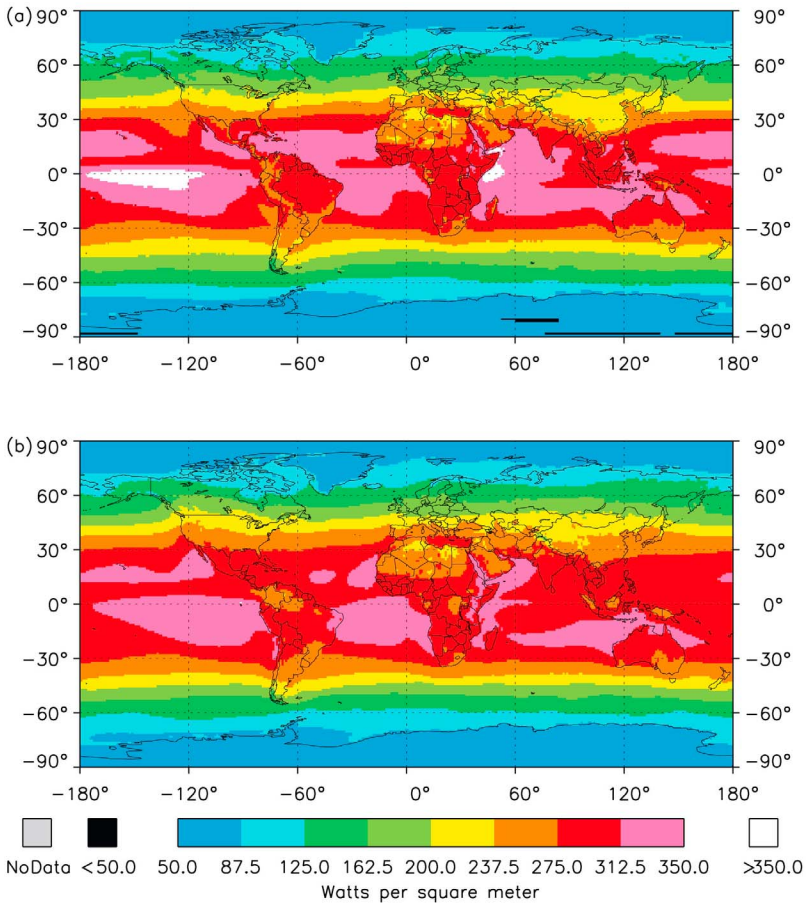

Figure 1. Annual average of ASR (W m^{-2}) for (a) EBAF and (b) GEOS-5 model.

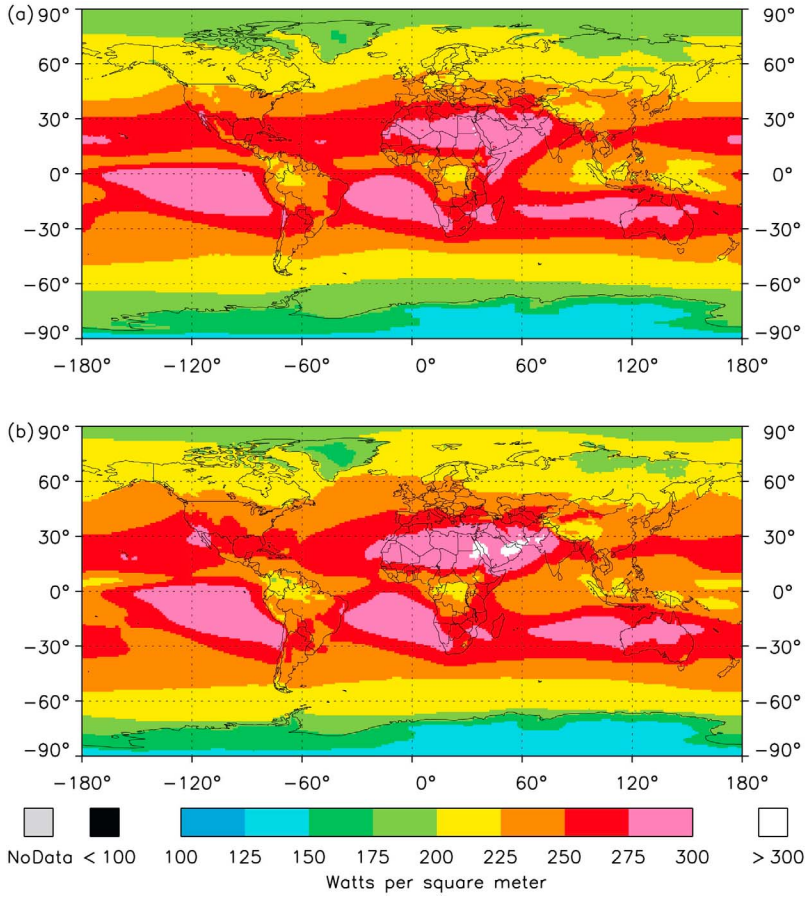

Figure 2. Annual average of OLR (W m^{-2}) for (a) EBAF and (b) GEOS-5 model.

applied to several data sets in order to compare them quantitatively and objectively and has application to the Coupled Model Intercomparison Project [*Potter et al.*, 2011].

[6] In this paper, the two data sets are described then the comparison method is applied to them. The method begins with the comparison of the annual mean maps of ASR and OLR. Next the annual mean is subtracted from the monthly mean maps to give the seasonal cycles of ASR and OLR. The time variations of these cycles are described in terms of principal components (PCs), and the spatial distributions are described by the corresponding empirical orthogonal components (EOFs). Measures are defined for the differences between the annual cycles of the two data sets. Inner products of the PCs and EOFs provide measures of the agreement or disagreement of the two data sets.

2. Data Sets

[7] The CERES Energy Balanced and Filled (EBAF) Edition 2.6 product [*Loeb et al.*, 2009] was used for this study. EBAF Ed2.6 contains monthly mean values of reflected shortwave and OLR at the top of atmosphere (TOA) from March 2000 through December 2010 on a 1° × 1° equal-angle grid. ASR is calculated as insolation minus the reflected shortwave. EBAF Ed2.6 is the net balanced CERES product and employs the latest instrument gains and spectral response functions. The EBAF product uses improved angular distribution models (ADMs) to convert the radiances into fluxes. The ADMs are selected based on scene type and viewing geometry [*Loeb et al.*, 2005, 2007], and MODIS cloud properties [*Minnis et al.*, 2011] help identify the scene types. Geostationary satellite flux data supplements the CERES data for computing the diurnal cycles in the fluxes, and in EBAF Ed2.6, scene dependent diurnal correction factors have been applied to the shortwave fluxes to remove geostationary satellite artifacts. To achieve the balance in the EBAF product, the shortwave and longwave fluxes of the CERES gridded monthly mean data product (SRBAVG-GEO) were adjusted within their ranges of uncertainty by use of an objective constrainment algorithm to remove the inconsistency between average global net TOA flux and heat storage in the earth–atmosphere system [*Loeb et al.*, 2009]. The data set and information regarding its production are available at http://ceres.larc.nasa.gov.

[8] Data from the GEOS-5 model (version Fortuna-2_2) [*Rienecker et al.*, 2008] were provided by the Global Modeling and Assimilation Office of Goddard Space Flight Center. The data set includes monthly mean ASR and OLR values from September 1987 through August 2007, for 20 years. Data are given on a grid with resolution of 1° in latitude and 1.25° in longitude and then were interpolated onto the CERES grid. In the model, monthly sea surface temperatures (SST) and sea ice were prescribed. The model was initialized with the atmosphere at the beginning of the run in 1987, but no atmospheric data were assimilated into

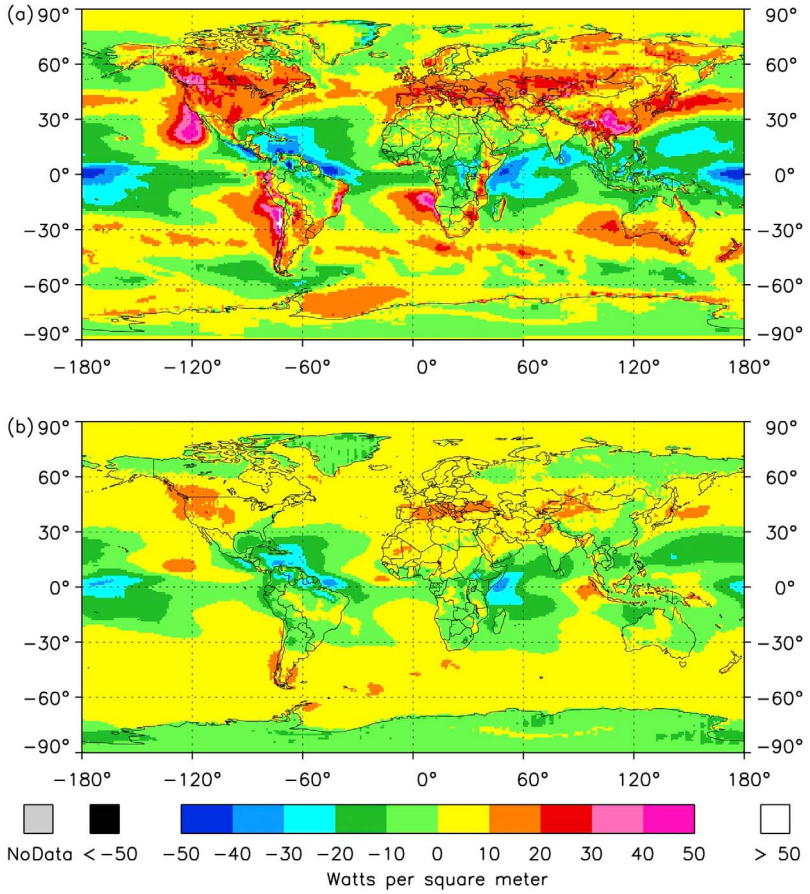

Figure 3. Difference between EBAF and GEOS-5 model annual average fluxes (a) ASR and (b) OLR.

the model after the initialization. Any effects of the initial conditions have vanished in the time before March 2000, when the EBAF data begins, so the only forcing is the prescribed varying SST and sea ice. The model includes moist processes, radiation, turbulent mixing, and surface processes. Clouds are generated by the model. ASR and OLR are strongly influenced by clouds, so that the differences between the model and EBAF depend to a large extent on clouds.

3. Comparison of Annual Means

[9] The first step in comparing the radiation fluxes is to examine the annual means of ASR and OLR and their differences, both globally and regionally. The seasonal cycles of ASR and OLR are then compared in section 4. The climatological mean, or canonical mean, for each calendar month was computed for the data period March 2000 through August 2007, the period of overlap between the two data sets. The annual mean was then computed as the average of the canonical monthly values. The seasonal cycle is taken to be the difference of the canonical months from the annual mean.

[10] Table 1 summarizes the global averages of the annual-mean ASR and OLR fluxes for the CERES-derived EBAF product and the GEOS-5 model results and their

Table 2. Root-Mean Square of the Seasonal Cycles of ASR and OLR for EBAF and GEOS-5 Model Over Land and Ocean

	RMS (W m^{-2})			
	ASR		OLR	
	EBAF	Model	EBAF	Model
Land	72.1	76.9	21.4	23.6
Ocean	73.3	74.1	12.1	13.7

Table 3. The First Three Normalized Eigenvalues, or Fraction of Variance Accounted for by the PCs, of the Seasonal Cycle of ASR and OLR

	ASR		OLR	
	EBAF	Model	EBAF	Model
Land				
1	0.957	0.956	0.880	0.874
2	0.029	0.026	0.054	0.059
3	0.009	0.011	0.032	0.033
Sum	0.995	0.993	0.966	0.966
Ocean				
1	0.961	0.963	0.783	0.706
2	0.024	0.020	0.084	0.130
3	0.010	0.012	0.047	0.064
Sum	0.995	0.995	0.914	0.900

Table 4. Root-Mean Squares for First Three Principal Components of the Seasonal Cycle of ASR and OLR for Land and Ocean for EBAF (R_n) and for GEOS-5 Model Results (r_n) (W m^{-2})

	ASR		OLR	
	EBAF	Model	EBAF	Model
Land				
1	70.6	75.2	20.1	22.0
2	12.2	12.5	4.9	5.7
3	6.7	7.9	3.8	4.3
Ocean				
1	71.8	72.7	10.7	11.5
2	11.3	10.4	3.5	4.9
3	7.2	8.1	2.6	3.4

differences. For the model, the OLR is 1.5 W m^{-2} greater than the ASR. This inequality is a result of the specified model ocean temperature. Thus, any amount of heat can be transferred to or from the model ocean without causing any change of the ocean temperature. Figure 1 shows the annual mean ASR for EBAF and for the model. The two data sets appear to agree well in the extratropics, but near the equator EBAF shows an annual mean ASR of more than 350 W m^{-2} over the central Pacific Ocean and over the western Indian Ocean, higher than for GEOS-5. Figure 2 shows the annual mean OLR maps for EBAF and for the model. The two maps appear to agree very well with subtle differences.

[11] The difference between the annual-mean maps is computed (model-EBAF), and this difference map is adjusted to have a zero global average by subtracting the global average difference from each region. Figure 3a shows the GEOS-5 model ASR minus the EBAF ASR minus the global-average difference between the model and EBAF (1.13 W m^{-2}, noted in Table 1). The largest positive differences are at the eastern part of the Pacific Ocean by North America and South America and at the eastern part of the Atlantic Ocean by southern Africa. These regions are marked by stratocumulus clouds. At about 45°N there is a band around the Earth of differences greater than 10 W m^{-2}, with larger values over the mountain states of the U.S. and the Himalayas, continuing into China. This band appears over North Africa and Western Europe, but it does not go over the Mediterranean Sea. There are regions of large negative differences over the tropical oceans, especially near the western parts of the Indian and Atlantic Oceans and over the central and western equatorial Pacific. These differences are due to the tropical clouds being too bright in the model [*Molod et al.*, 2012].

[12] Figure 3b shows the model OLR minus the EBAF OLR minus the global-average difference between the model and EBAF (3.13 W m^{-2}, Table 1). Over the Southern Ocean the agreement is excellent. Over the Pacific and Atlantic Oceans outside the tropics, the agreement is again very good. The band of differences in ASR which was noted around the Earth near 45°N appears only as a high over the western U.S., the Mediterranean Sea, and Central Asia. In the OLR over western tropical oceans there are negative differences with magnitudes greater than 10 W m^{-2}, which match the pattern of ASR negative differences, although

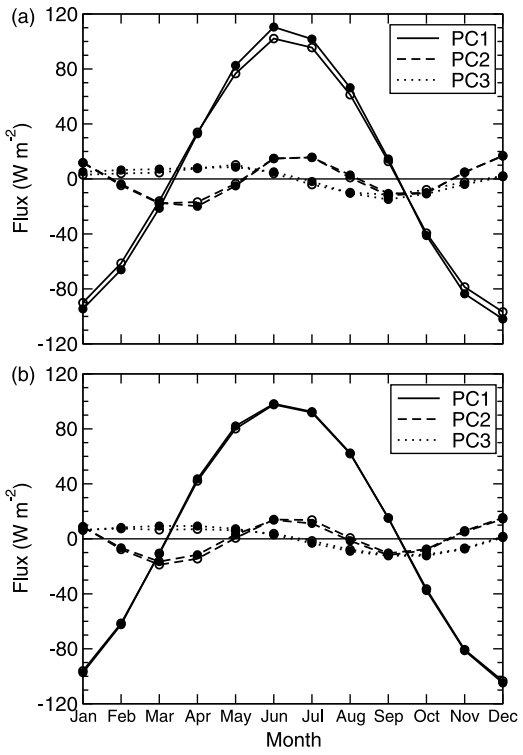

Figure 4. First three PCs for the annual cycle of ASR for EBAF (open circle) and the GEOS-5 model (solid circle) (a) over land and (b) over ocean (W m^{-2}).

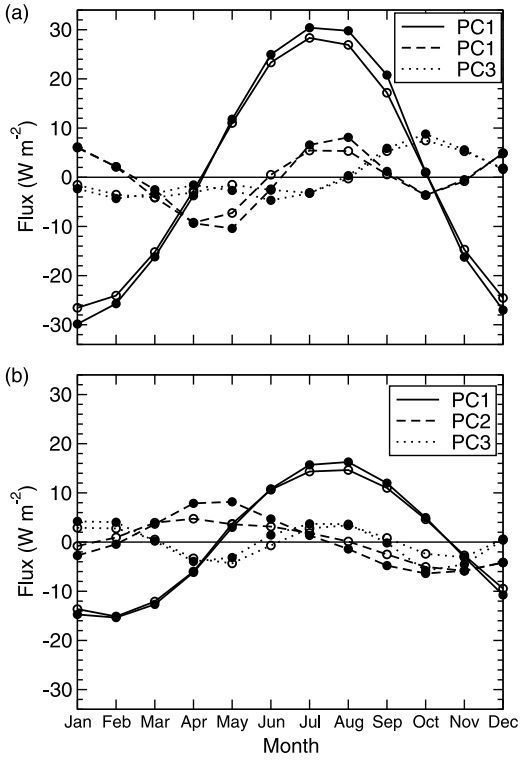

Figure 5. First three PCs for the annual cycle of OLR for EBAF (open circle) and the GEOS-5 model (solid circle) (a) over land and (b) over ocean (W m^{-2}).

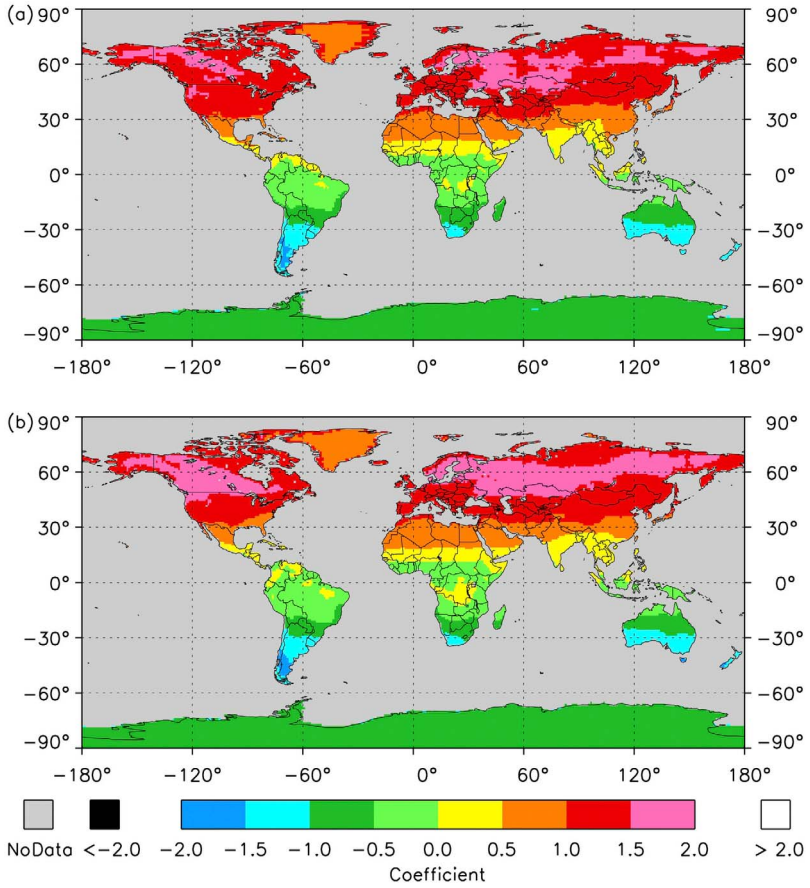

Figure 6. The first EOF of annual cycle of ASR over land for (a) EBAF and (b) GEOS-5 model.

they are not as strong in OLR as for ASR over the subtropical regions. One possible explanation for these differences over the oceans is there may be low clouds present in the EBAF data that are not present in the model results. The absence of clouds in the model would give higher ASR, and with low clouds there would be little effect on the OLR. The negative differences in both ASR and OLR may illustrate the model's systematic excessive shortwave and longwave cloud forcing as a result of its tendency to place deep convection on the edge of warm tropical SSTs as opposed to being more centered over the western Pacific Warm Pool. The positive OLR differences just west of Indonesia along the equator are in a region where convection in GEOS-5 is systematically missing. Uncovering and confirming the cause of the annual mean differences is beyond the scope of this paper and will need to be addressed in a future study.

[13] The question arises, how much do the model results differ from the EBAF data set overall? In other words, can one provide a number that summarizes the differences between the model and EBAF maps for ASR and a number for the OLR differences? The differences between the model and EBAF maps of ASR or OLR may be measured as the root-mean square (RMS) of the differences over the Earth. The RMS of a function $f(x)$ defined in the space domain is its absolute value and is computed as

$$\mathrm{RMS}\,[f(x)] = \left[\sum_{x=1}^{N} w_x f^2(x)\right]^{1/2}, \qquad (1)$$

where x is an index for the grid boxes covering the Earth, the w_x are the area weights for the grid boxes, and $\sum_{x=1}^{N} w_x = 1$. An important property of RMS as a measure of the differences is that it takes into account all points of the map, whereas a measure such as range may be determined by merely two outliers. Table 1 lists the RMS of the difference maps of ASR and OLR in Figure 3. Although locally the model and EBAF ASR values differ by as much as 50 W m^{-2}, the RMS is about 14 W m^{-2}, and likewise the RMS of the OLR differences is about 8 W m^{-2}.

4. Variations of the Seasonal Cycles

[14] The seasonal cycles of ASR are due to the seasonal cycles of solar declination and Earth-Sun distance and the variations of cloud amount and snow cover. These ASR cycles create cycles of atmospheric and oceanic temperature with resulting cycles of circulation, cloud amounts, and snow cover in a coupled system with multiple feedbacks. The seasonal cycles of ASR and OLR are represented here with one-month resolution. There are 12 maps of monthly means of ASR and OLR that define the time-varying fields on a grid with 64800 grid points. For the model field, the annual mean map is subtracted from the map for each month to give the model seasonal cycle, which is denoted as $y(x, t)$. Likewise the seasonal cycle of EBAF is formed and is denoted as $Y(x, t)$.

[15] How closely does the time response of $y(x, t)$ match that of $Y(x, t)$? This section describes the method for answering

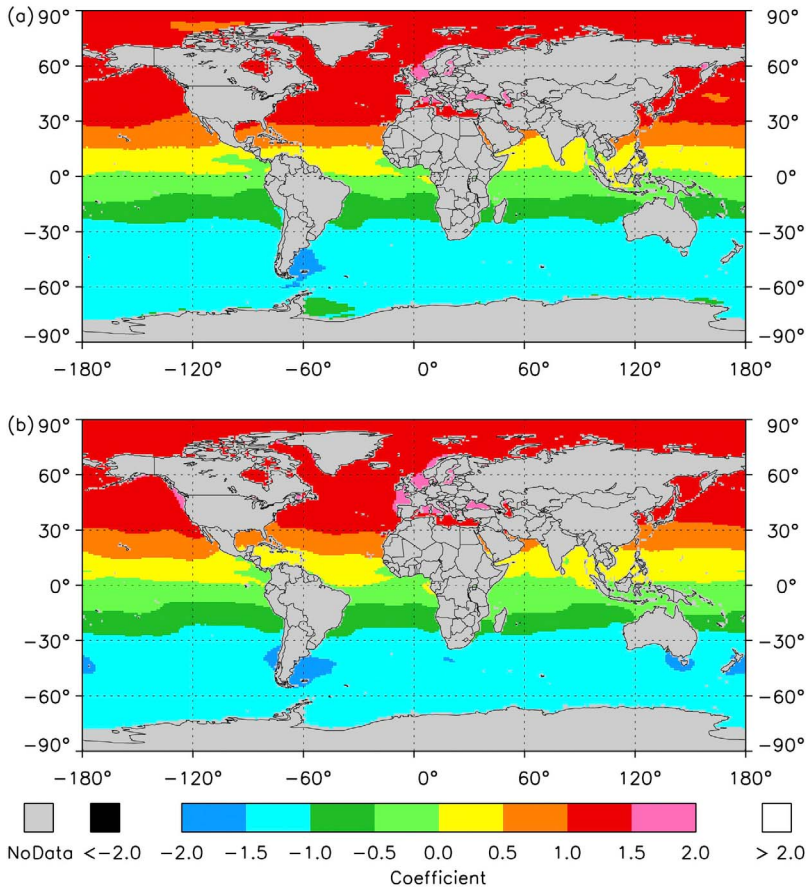

Figure 7. The first EOF of annual cycle of ASR over ocean for (a) EBAF and (b) GEOS-5 model.

this question. Principal components are computed for the time variations of the EBAF product and for the model for the ASR and OLR. Empirical orthogonal functions EOFs computed with each of the PCs describe the associated spatial variations. The PCs of the EBAF product are used as a basis set for the differences between the model and the EBAF seasonal cycles. The RMS of this difference is a measure of the disagreement between the two data sets. The agreement of the time responses of the two data sets is expressed by the matrix of correlation coefficients of the PCs of the two sets, and the agreement of the spatial variability between the data sets is expressed by the correlation coefficient matrix of the EOFs of the two sets. Relations among these correlation matrices are presented in Appendix A.

4.1. Magnitude of the Seasonal Cycles

[16] For the seasonal cycle we deal with a function f of time t, where time $t \in [1, 12]$. The RMS for a monthly mean is then given by

$$\mathrm{RMS}\,[f(t)] = \left[\sum_{t=1}^{12} f^2(t)/12\right]^{1/2}. \qquad (2)$$

If the field is a function of both space and time, the operations of equations (1) and (2) are combined to give RMS by

$$\mathrm{RMS}\,[f(x,t)] = \left[\sum_{t=1}^{12}\sum_{x=1}^{N} w_x f^2(x,t)/12\right]^{1/2}. \qquad (3)$$

The immense heat capacity of the ocean causes a lag of temperature, and hence OLR, relative to that of land. By separating land and ocean in the PC analysis, the annual cycle is better represented for each surface type. A land/ocean mask partitions the Earth's surface into 21259 land grid points and 39734 ocean grid points. A third category, coast, accounts for 3807 grid points and is omitted in this study.

[17] The RMSs of the seasonal cycles of ASR and OLR [from equation (3)] are listed in Table 2; the RMS values provide a measure of the size of the cycles. The seasonal cycle of ASR is more than three times that of OLR. *Mlynczak et al.* [2011] point out that the seasonal cycle of OLR is smaller than that for ASR because of storage and transport of heat. The RMSs of ASR for land and ocean for the model and for EBAF agree extremely well, and the RMSs for OLR agree quite well also. However, agreement of these numbers by themselves does not assure that the time and space variation of the two data sets are similar.

4.2. Principal Component Analysis

[18] Additional information is required to diagnose where and when differences occur. A comparison of the two data sets should take the time variation into account rather than comparing each monthly mean map individually. Principal component analysis simplifies this task by reducing to a minimum the number of terms required to describe the data sets. *Mlynczak et al.* [2011] show that, using principal component analysis, the EBAF ASR time variation over

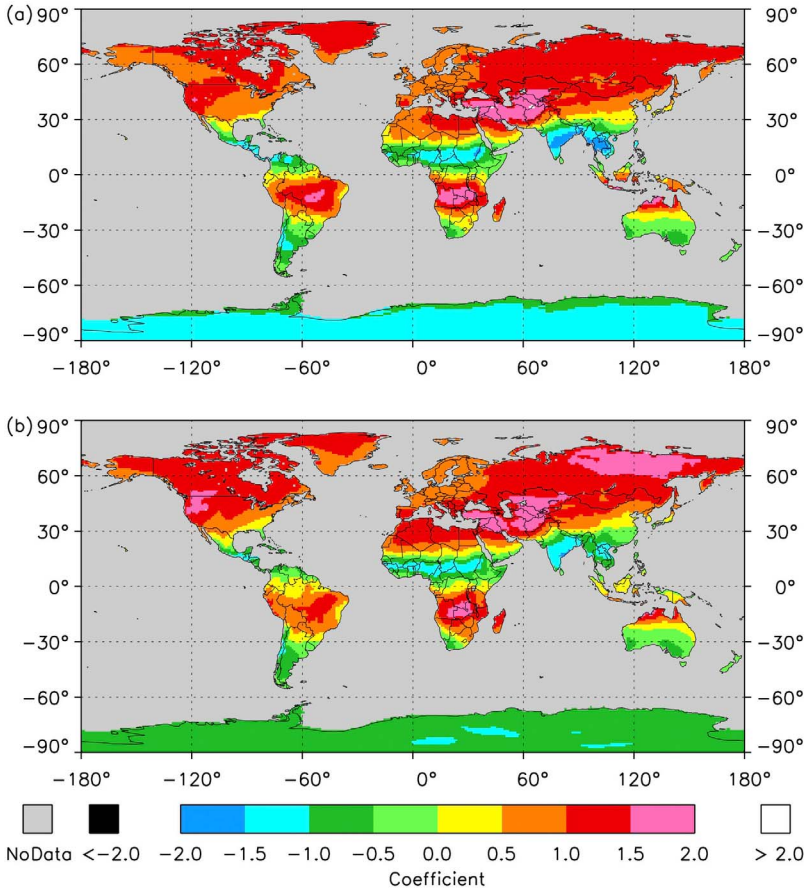

Figure 8. The first EOF of annual cycle of OLR over land for (a) EBAF and (b) GEOS-5 model.

land can be described with one term which accounts for 95.7% of the total variance, or the area-weighted sum of squares over the time and space domain. That study showed that the first principal components provided the most information about the ASR and OLR variations for both land and ocean.

[19] The computation of the PCs and EOFs shown in this study is described by *Mlynczak et al.* [2011]. The PCs, or time variations, are denoted with $\Phi_n(t)$, where $n \in [1, 12]$. The associated spatial variations are the EOFs, denoted by $\Psi_n(x)$. The EBAF seasonal cycle $Y(x, t)$ can be expressed as

$$Y(x,t) = \sum_{n=1}^{12} R_n \Phi_n(t) \Psi_n(x). \quad (4)$$

R_n is the RMS for the nth PC term, i.e., the term magnitude. The principal components are orthogonal and are normalized such that

$$\sum_{t=1}^{12} \Phi_m(t) \Phi_n(t)/12 = \delta_{mn}. \quad (5)$$

The GEOS-5 model seasonal cycle likewise has normalized PCs $\varphi_n(t)$ and EOFs $\psi_n(x)$ and can be written as

$$y(x,t) = \sum_{n=1}^{12} r_n \varphi_n(t) \psi_n(x), \quad (6)$$

where r_n is the RMS for the model nth PC term. The computation of R_n and r_n will be discussed in the next section.

[20] Comparison of equations (4) and (6) shows that there are three requirements for the two data sets to agree. First, the RMS accounted for by each term must be the same. Second, both sets of principal components, which describe the time variations, must be the same. Finally, the empirical orthogonal functions, which describe the spatial variations, must be the same. An exception to this can occur if two eigenvalues are equal in one set, resulting in a degenerate case in which the PCs are not uniquely defined [*North et al.*, 1982]. The comparison of each of these factors is considered next in order to assess how well the data sets agree and how they differ in terms of time and space.

4.3. Comparison of Term Magnitudes

[21] The eigenvalues, Λ_n for EBAF and λ_n for the model, are normalized such that their sum is one. Each eigenvalue is the fraction of variance, or fraction of root-sum-square, associated with its PC. The first three normalized eigenvalues for the EBAF and for the model results are listed in Table 3 for land and ocean. In each case the first term accounts for the vast majority of the variance. Three terms account for over 99% of the variance for ASR for both land and ocean. The eigenvalues for the two data sets agree very well for all but the ocean OLR case. The ocean OLR eigenvalues may differ from the others because the RMS of the ocean OLR (Table 2) is small by comparison and thus sensitive to small changes, which would cause more variance to appear in the higher-order terms. Therefore the ocean

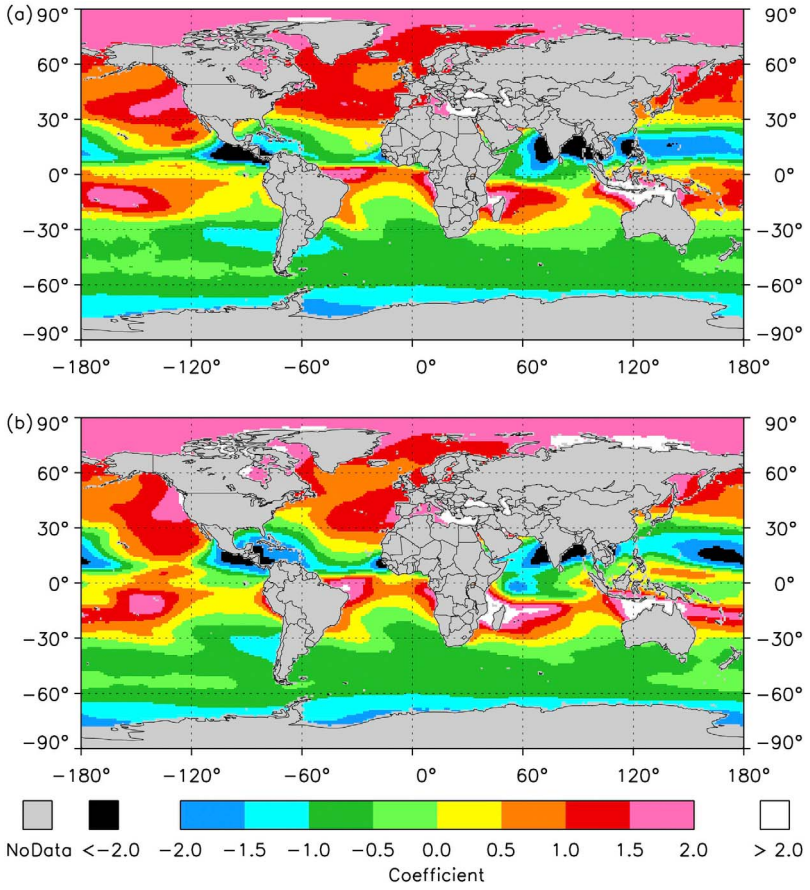

Figure 9. The first EOF of annual cycle of OLR over ocean for (a) EBAF and (b) GEOS-5 model.

OLR eigenvalues do not decrease as rapidly as for the other cases. Because the first term accounts for the vast majority of the time and space variation, principal component analysis provides a tool for simplifying the examination of the differences between the two data sets.

[22] The term magnitudes R_n and r_n for each PC_n are computed from the seasonal cycle RMS and the eigenvalues: $R_n = \text{RMS} \times \Lambda_n^{1/2}$ and $r_n = \text{RMS} \times \lambda_n^{1/2}$. The first three are listed in Table 4 for land and ocean for EBAF and for model results for ASR and OLR. The term RMSs agree within 1 W m^{-2} except for PC_2 for OLR over ocean, so the magnitudes are close. The next step is to examine the PCs and EOFs and their differences qualitatively, then quantitatively.

4.4. Qualitative Comparison of Time Responses

[23] The two sets of normalized principal components $\Phi_n(t)$ and $\varphi_n(t)$ define the time responses for the two sets of data. We examine the first three PCs, which account for the vast majority of variance of the seasonal cycle. The dimensionalized PCs $R_n\Phi_n(t)$ and $r_n\varphi_n(t)$ are used in order to keep the physical magnitudes of the terms evident. Figure 4 shows the first three dimensionalized principal components for ASR over land and ocean. OLR is the flow of energy from the system, and Figure 5 shows the first three dimensionalized principal components for the seasonal cycle of OLR for land and for ocean.

[24] The close similarity of the PCs shows that the time responses of the EBAF data set match those of the model. For example, Table 4 shows that for ASR over land, R_1 = 70.6 W m^{-2} and r_1 = 75.2 W m^{-2}. Consequently, the dimensionalized PC_1s in Figure 4a differ slightly in magnitude but have the same shape. The R_1 and r_1 values are much closer for ASR over ocean; their PCs in Figure 4b are extremely close. When the dimensionalized PCs are divided by R_n and r_n, respectively, the $\Phi_n(t)$ and $\varphi_n(t)$ are very nearly identical for n = 1, 2, and 3. The relationships between $\Phi_n(t)$ and $\varphi_n(t)$ will be discussed further in section 4.8.

[25] For ASR and OLR, PC_1 of both EBAF and the model is an annual cycle with a shape close to a sine wave, and PC_2 is a wave 2, or semiannual cycle, in each case except that of OLR over ocean, for which PC_2 is an out-of-phase annual cycle, and PC_3 is a semiannual cycle. PC_3 is an out-of-phase annual cycle for ASR and for OLR over land. In each case the model response has the same shape and phase as the EBAF result for the first three PCs. PC_1 for ASR is strongly influenced by orbital mechanics, but PC_3 is the out-of-phase response to PC_1 and describes snow cover and clouds lagging PC_1 as well as monsoonal variations. The ASR semiannual cycle describes the polar day and night [*Mlynczak et al.*, 2011]. For OLR, PC_1 is the annual cycle responding directly to ASR PC_1. The PC_2 over land and PC_3 over ocean are semiannual cycles that describe the variations over the monsoon regions and the movements of the ITCZ. The PC_3 over land and PC_2 over ocean give the out-of-phase response of the OLR to ASR PC_1.

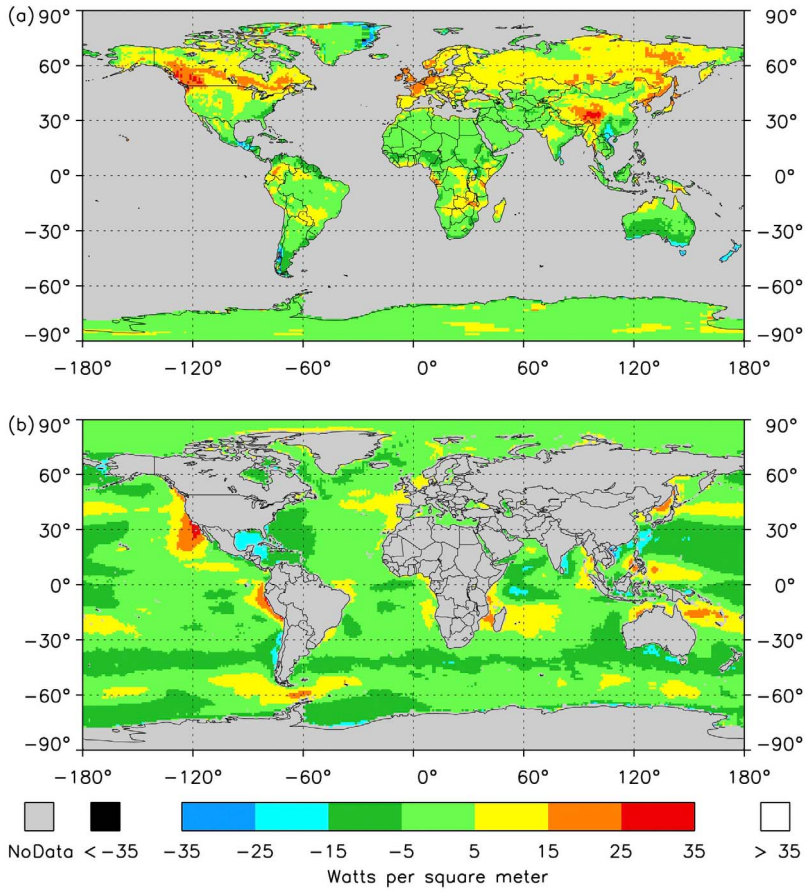

Figure 10. Δ_1 maps for the annual cycle of ASR over (a) land and (b) ocean (W m^{-2}).

4.5. Qualitative Comparison of Spatial Variations

[26] The question remains as to how well the spatial variations, defined by the EOFs, agree. Figure 6 shows EOF_1 of the seasonal cycle of ASR over land for EBAF and for the model. The two maps agree over most of the land. Figure 7 shows the EOF_1 for ASR over ocean. Subjectively the agreement between EBAF and the model appears to be very good. These results show that the energy entering the system, governed by orbital mechanics and clouds, appears to be accounted for accurately in the model. The map of the annual cycle of the energy leaving the system is given by EOF_1 of OLR, which is shown by Figure 8 for land. The major differences are over Canada and Siberia, where the model computes larger values than EBAF has. Over Antarctica the model shows a smaller cycle than EBAF, and there are small differences over equatorial South America. Figure 9 shows EOF_1 for OLR over ocean. The annual cycle over the Pacific Intertropical Convergence Zone is larger for the model than for EBAF. The observed sea-surface temperature is used by the GEOS-5 model, so this OLR difference would indicate a greater variation of cloud forcing over this region in the model than in EBAF. This effect could be attributed to differences of cloud fraction or of cloud top height. Over the North Atlantic Ocean south of Greenland the OLR variation is less in the model than in EBAF.

4.6. Quantitative Comparison of Spatial Variations

[27] In order to compare the spatial variations of the model with EBAF, a single time basis should be used. The normalized EBAF and model PCs, $\Phi_n(t)$ and $\varphi_n(t)$ respectively, differ in principle, so that there are two different descriptions of the time variation. We select the EBAF PCs $\Phi_n(t)$ as the basis for time description and project the difference $y(x, t) - Y(x, t)$ onto the $\Phi_n(t)$, giving a map of differences for each order of the EBAF time response:

$$\Delta_n(x) = \sum_{t=1}^{12} \Phi_n(t)[y(x,t) - Y(x,t)]/12. \tag{7}$$

Note that the projection of $Y(x, t)$ onto $\Phi_n(t)$ is the EOF $\Psi_n(x)$. The difference $y(x, t) - Y(x, t)$ is then the sum of the contributions of the Δ_n with each time response:

$$y(x,t) - Y(x,t) = \sum_{n=1}^{12} \Delta_n(x)\Phi_n(t). \tag{8}$$

Each location has a set of $\Delta_n(x)$ values, so that one can develop a map for each Δ_n term.

[28] For ASR, the major part of the time response of $Y(x, t)$ is given by $\Phi_1(t)$, which is seen to represent the annual cycle, so examination of the maps of Δ_1 provides information about most of the differences between the two data sets for the annual cycle. Therefore, we can examine one Δ_1 map for

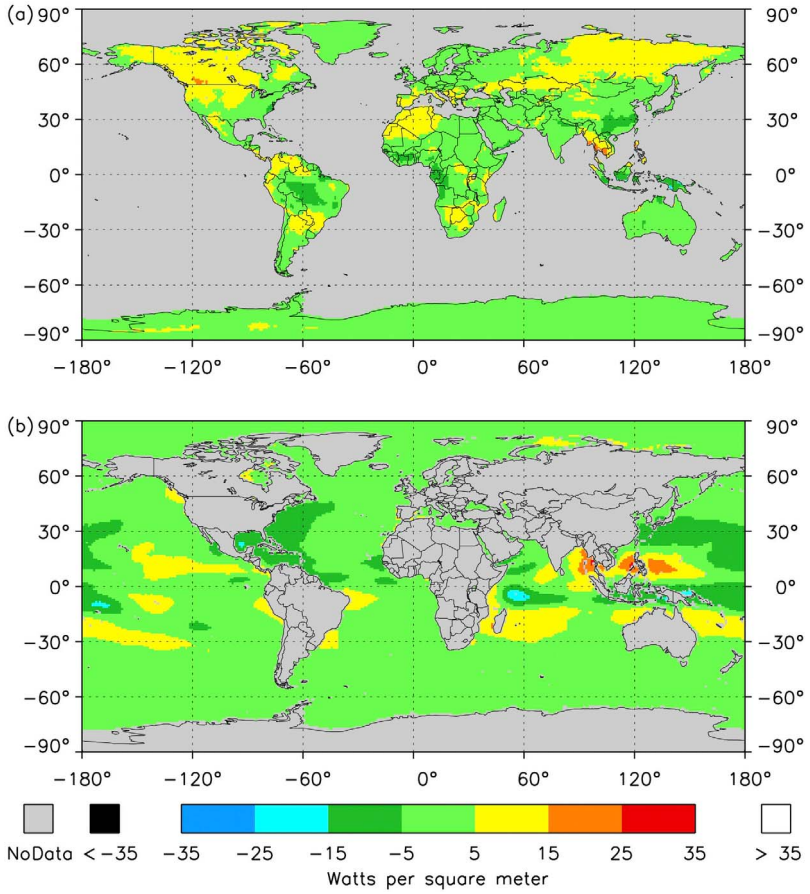

Figure 11. Δ_1 maps for the annual cycle of OLR over (a) land and (b) ocean (W m^{-2}).

ASR rather than 12 pairs of monthly mean maps of ASR in order to compare the model with EBAF results. Figure 10 shows the Δ_1 maps for ASR over land and over ocean. These maps show the difference between EBAF and the model results for the annual cycle of ASR. Higher-order Δ_n maps show other components of the seasonal cycles, e.g., the semiannual cycles. As normalized, $\Phi_1(t)$ of ASR is positive in July and negative in January (see Figure 4 for shape of normalized PC), so that in the Northern Hemisphere a positive Δ_1 indicates a positive difference in NH summer, that is, the model value is greater than the EBAF value when $\Phi_1(t)$ is positive. In the Southern Hemisphere a negative Δ_1 indicates a positive difference in SH summer.

[29] For the Δ_1 map for ASR over land (Figure 10a), many regions agree with differences less than 5 W m^{-2}. Over Canada and the Himalayas, the model is 15 to 25 W m^{-2} greater than the EBAF results. Over ocean (Figure 10b), the ASR for most regions agrees within 5 W m^{-2}, but there are regions where the differences are greater than 15 W m^{-2}. Moreover, these regions with differences greater than 15 W m^{-2} are spatially coherent over hundreds of kilometers, indicating that they are not simply scattered errors. Over the eastern Pacific Ocean off the coast of North America there is an area where the annual cycle of the model is 15 to 35 W m^{-2} higher than the EBAF annual cycle. Because of the reversal of sign between the Northern and Southern Hemispheres for the annual cycle, this corresponds to the region over the Pacific Ocean west of Chile, where Δ_1 has a value between −15 and −25 W m^{-2}.

[30] Figure 11 shows the Δ_1 maps for OLR over land and over ocean. There are areas of positive differences from 5 to 15 W m^{-2} over North America and Siberia, which are attributed to clouds having less effect in the model than in EBAF data. This OLR difference together with the ASR difference over these regions indicates that the annual cycle of clouds is not sufficiently strong. Over most of the ocean the model and EBAF agree within 5 W m^{-2}. Over the eastern Indian Ocean and north of the Maritime Continent, the OLR difference is 15 to 25 W m^{-2}.

[31] The application of the $\Delta_n(x)$ is illustrated by examining the time variations of ASR over three regions for which $y(x, t) - Y(x, t)$ is large and their reconstruction using one and two terms (equation (8)) can be demonstrated. The first is the region over the eastern Pacific Ocean off the coast of North America, noted earlier. Figure 12a shows the yearly variation of ASR averaged over this region, from 27°N to 30°N and 119°W to 123°W, for both the model and EBAF. The clear-sky ASRs from the model and from EBAF for this region are nearly identical, so the difference is due to the difference of the cloud forcing of the ASR. The model ASR is much larger than EBAF indicates for April through August, indicating insufficient cloud forcing over this region in the model. This could be due to errors in cloud amount or cloud optical properties. Figure 12b shows the seasonal

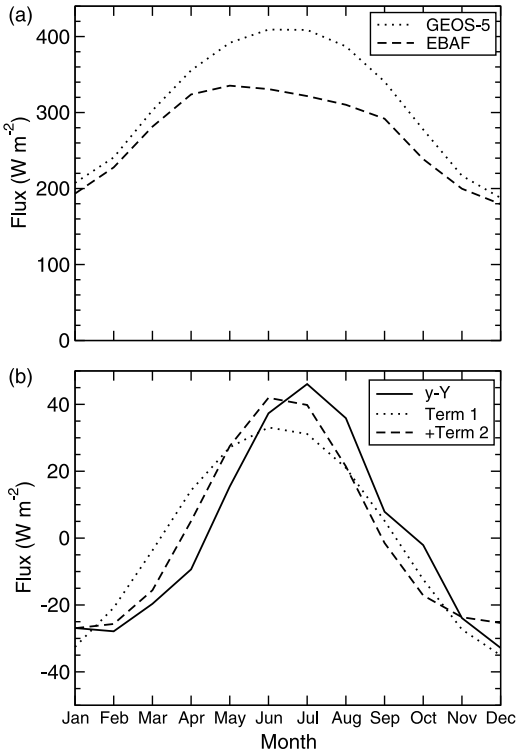

Figure 12. (a) Absorbed solar radiation flux over year for GEOS-5 model and EBAF and (b) the difference in the seasonal cycles $y(x, t) - Y(x, t)$ and its expression in terms of $\Delta_1(x)$ and $\Delta_2(x)$ for region in the eastern Pacific Ocean.

cycle difference $y(x, t) - Y(x, t)$ and also the reconstruction of that difference using equation (8) with $\Delta_1 = 24.34$ and $\Delta_2 = 7.12$. Most of the difference between the two data sets is described by the first two Δ_n terms, but higher-order terms are needed to correct the phasing difference.

[32] A second ASR example is shown by Figure 13a for a region over the Gulf of Mexico from 23°N to 27°N and 86°W to 94°W. The ASR agrees very well from October through May, but from June through September, the ASR decreases markedly for the model. Again the clear-sky results are very close for the two data sets, so the seasonal cycle difference shown in Figure 13b is due to the cloud forcing of ASR. In this case the model computes the presence of clouds during summer which do not appear. The reconstruction of the difference using $\Delta_1 = -20.61$ and $\Delta_2 = -5.62$ shows that higher-order terms are required to describe the difference for this region.

[33] Finally, Figure 14a shows the ASR for a year averaged over a region in the Canadian Rockies from 50°N to 54°N and 116°W to 120°W. The model ASR is high compared to EBAF during most of the year. As with other cases, the clear-sky ASR flux is very close between the two data sets, so that the high ASR in summer indicates a paucity of clouds in the model results. Figure 14b shows the difference $y(x, t) - Y(x, t)$. The reconstruction using two terms with $\Delta_1 = 25.45$ and $\Delta_2 = 4.41$ describes the difference well.

4.7. Global Measure of Time-Resolved Maps

[34] The RMS of the difference between the GEOS-5 model and EBAF over the globe and for the seasonal cycle can be expressed in terms of the $\Delta_n(x)$. By use of the orthogonality of the $\Phi_n(t)$, equation (8) gives

$$\mathrm{RMS}\,[y(x,t) - Y(x,t)] = \left(\sum_{n=1}^{12} \{\mathrm{RMS}\,[\Delta_n(x)]\}^2\right)^{-1/2}. \quad (9)$$

The RMS$[\Delta_n(x)]$ is a global measure for the difference of the model field from the EBAF field for the time variation defined by $\Phi_n(t)$. The sum of the squares on the right hand side of equation (9) gives the sum of the squares of the difference of the two fields over the globe and year. The RMS of the difference between the model and EBAF fields is listed in Table 5 for the four cases: ASR and OLR for land and ocean. The RMSs of the seasonal cycles of EBAF, from Table 2, are also listed for comparison.

[35] It is useful to examine the effect of the order n on the magnitude of the Δ_n. Figure 15 shows the dependence of the RMS of the Δ_n on the order n for each of the four cases. The RMS of the EBAF term is also shown for comparison. The RMS of Δ_1 for ASR over land is 8 W m^{-2}, and the RMS of the term itself is 71 W m^{-2}, showing that the model is accounting for most of the variance globally, even though some regions differ strongly. The Δ_1 is by far the largest term for each case, so that higher-order terms are not considered here. This is another justification for using PCs as the basis set for comparing data sets. The power in the terms of order three and higher of both EBAF and the model data sets decreases rapidly with order. The higher-order terms may be thought of as being effectively random, so the Δ_n have a larger RMS than do the EBAF terms.

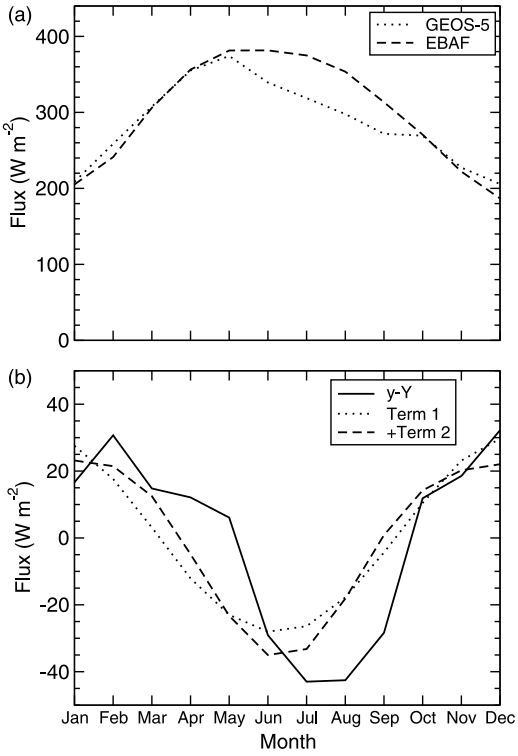

Figure 13. (a) Absorbed solar radiation flux over year for GEOS-5 model and EBAF and (b) the difference in the seasonal cycles $y(x, t) - Y(x, t)$ and its expression in terms of $\Delta_1(x)$ and $\Delta_2(x)$ for region in the Gulf of Mexico.

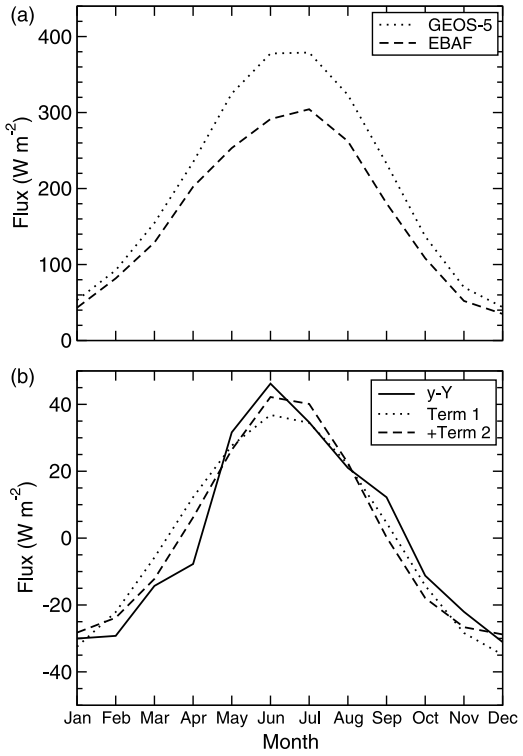

Figure 14. (a) Absorbed solar radiation flux over year for GEOS-5 model and EBAF and (b) the difference in the seasonal cycles $y(x, t) - Y(x, t)$ and its expression in terms of $\Delta_1(x)$ and $\Delta_2(x)$ for region in the Canadian Rockies.

4.8. Quantitative Comparison of Principal Components and Empirical Orthogonal Functions

[36] The principal components describe the time behavior of the system. One important question is how well do the principal components of the two data sets agree? The quantitative comparison of the sets of PCs is now considered. The correlation of the normalized PCs $\varphi_n(t)$ with $\Phi_n(t)$ is

$$\mathbf{H}_{mn} = \sum_{t=1}^{12} \varphi_m(t)\Phi_n(t)/12. \tag{10}$$

These correlations form the correlation matrix relating the two sets of normalized PCs. If and only if the two sets agree perfectly, $\mathbf{H}_{mm} = 1$ and $\mathbf{H}_{mn} = 0$ for $m \neq n$, i.e., $\mathbf{H}$ is the identity matrix. The diagonal terms are of primary interest and are listed in Table 6. For ASR $\mathbf{H}_{11}$ is 1.000 to three decimal places, and for OLR it is 0.999, indicating excellent agreement for $\varphi_1(t)$ and $\Phi_1(t)$. The similarity of the dimensionalized PCs shown in Figures 4 and 5 was noted in section 4.4. The higher-order $\mathbf{H}_{mm}$ decrease slowly, showing good agreement even to the fourth order. For $n = 1$, 2 and 3 the off-diagonal terms are small, so that

$$\Phi_n(t) \approx \varphi_n(t) \text{ for } n = 1, 2 \text{ and } 3. \tag{11}$$

Table 3 shows that these first three terms for ASR over land account for 99.5% of the variance for EBAF and 99.3% for the model results. The eigenvalues for orders greater than 3 have very little power associated with them, so these terms are in the noise level. Thus the time responses of the two data sets agree very well.

[37] The empirical orthogonal components describe the spatial behavior of the system, analogous to the PCs' description of the time behavior. One can similarly define a transformation matrix that relates the EOFs of the two data sets, enabling the quantification of the agreement of the spatial behavior of the two data sets in a manner similar to that for the PCs. The correlation of the model EOF_m with EBAF EOF_n is

$$\mathbf{K}_{mn} = \sum_x w_x \psi_m(x)\Psi_n(x). \tag{12}$$

The same considerations apply to the $\mathbf{K}$ matrix as for the $\mathbf{H}$ matrix. The first four diagonal terms of $\mathbf{K}$ are listed in Table 7. For first order, the ASR agrees very well, but the OLR does not agree so well, especially for ocean. The

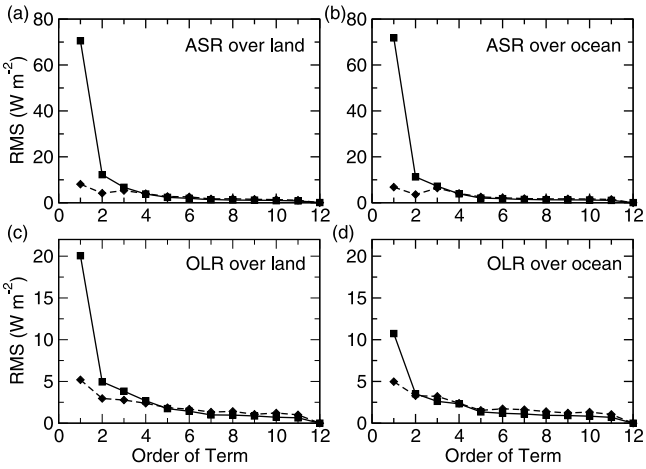

Figure 15. Comparison of EBAF term RMS (square symbols with solid lines) with Δ_n term RMS (diamond symbol with dashed line) as a function of order of term for (a) ASR over land, (b) ASR over ocean, (c) OLR over land, and (d) OLR over ocean (W m^{-2}).

Table 5. Global Mean Annual-Average RMS of Difference in Seasonal Cycles Between the GEOS-5 Model and EBAF (W m^{-2})

Quantity	RMS of Difference	RMS of EBAF
ASR over Land	12.3	72.1
ASR over Ocean	11.9	73.3
OLR over Land	7.9	21.4
OLR over Ocean	8.1	12.1

Table 6. Diagonal Terms of H_{mn} Matrices for ASR and OLR Over Land and Over Ocean

	ASR		OLR	
	Land	Ocean	Land	Ocean
1	1.000	1.000	0.999	0.999
2	0.995	0.990	0.967	0.940
3	0.973	0.991	0.976	0.922
4	0.966	0.976	0.932	0.926

Table 7. Diagonal Terms of K_{mn} Matrices for ASR and OLR Over Land and Over Ocean

	ASR		OLR	
	Land	Ocean	Land	Ocean
1	0.996	0.996	0.974	0.902
2	0.942	0.950	0.837	0.683
3	0.747	0.648	0.754	0.458
4	0.648	0.479	0.664	0.567

relations among the **H** and **K** matrices and the Δ_n are examined in Appendix A.

5. Discussion

[38] Absorbed solar radiation ASR flux and outgoing longwave radiation OLR flux from the CERES EBAF data set are compared with the ASR and OLR of the GEOS-5 data set. ASR and OLR at the top of the atmosphere computed from CERES measurements have a minimum of modeling and dependence on ancillary data, so these quantities are well-suited for comparing with model results. There are several levels of detail which one may address when comparing data sets which describe time-varying fields. The hierarchy of questions is summarized in Table 8.

[39] The exceptionally good agreement of the first principal components for ASR is noteworthy. The expression of the TOA annual cycle of insolation has been discussed by *Wilber et al.* [2006]. Although insolation at TOA may be computed to high accuracy, the ASR depends also on albedo, which is strongly affected by clouds. The difference of the model ASR annual cycle and the EBAF annual cycle was projected onto the first observed principal component to form Δ_1. For a few areas, differences were noted up to 35 W m^{-2}, but for most of the Earth, the differences were less than 5 W m^{-2}. The agreement of $\Phi_1(t)$ with $\varphi_1(t)$ over land for ASR, measured by H_{11}, was 1.000 to three decimal places. This agreement in the time variation may be fortuitous in consideration of these spatial differences, or it is possible that there is a reason for the agreement in the time domain.

[40] For the CERES EBAF data set, the error for ASR is about 1 W m^{-2} (less than 0.5%) and for OLR is 2.4 W m^{-2} (about 1%). There are three types of errors in the CERES data products: bias (or off-set) errors, gain errors, and errors due to spatial and temporal sampling and angular directional models. The bias and gain errors are instrument errors, and the remaining errors are incurred while generating higher-level data products. Bias and gain errors at the mean flux level appear in the annual-mean maps of ASR and OLR and are part of the global average, which was discussed in section 3. The effects of gain errors in the seasonal cycles are less than 1 W m^{-2}. Any errors due to sampling or angular directional models would not be present in the model results and would be seen in the Δ_n maps as patterns related to the orbit. Such error signatures do not appear in these maps, thus validating that they are less than 1 W m^{-2} in the EBAF data product.

6. Concluding Remarks

[41] A method for the comparison of two fields that vary in time and space is applied to compare the energy balanced EBAF data set based on observations of ASR and OLR from CERES with the results from the GEOS-5 model computations. The annual average maps of ASR and OLR are first examined, then the time variation is expressed in terms of principal components for the time dependence and their associated empirical orthogonal functions for the spatial distributions. It is shown that the first two principal components describe the annual and semiannual cycles for ASR and OLR and that the model and EBAF principal components are very similar.

[42] The quantitative comparison of EBAF with the model results are made at several levels of detail. The RMS differences are computed for the complete time and space domain to give a single number for ASR and for OLR. In order to characterize the agreement of time response of the model results with EBAF, correlation matrices are defined which describe the level of agreement between the principal components of the model and of EBAF. The parameters used here and the questions that they address are summarized in Table 8.

[43] The seasonal cycles of ASR and OLR can be expressed by two terms to a high accuracy, so that the principal components which are used to investigate the time and space variations have physical meanings as the annual and semiannual cycles. However, the study of a field that requires a number of principal components for its description may not represent real physical processes in a meaningful way.

[44] The differences of the two data sets over the year are projected onto the EBAF principal components to form a set of maps, Δ_n. The RMSs of Δ_n comprise a small set of numbers for goodness of fit for the seasonal cycles of the two data sets. If one wished to compare the seasonal cycles of parameters from many models with observations as was done for cloud forcing by *Cess et al.* [1997], RMS of Δ_n is a suitable measure of goodness of fit of the many models with the observed fields.

[45] Differences between the model results and EBAF have been noted. Most of the differences appear to be explained by the differences in clouds as generated by the model. The major difference between the ASR of the two data sets is that the annual cycle of the model has a greater magnitude than that of the EBAF. To diagnose the reasons for these differences and to improve the parameterizations by which the clouds are computed will require further work.

[46] We also anticipate that this analysis technique would be useful in the analysis of model simulations from the fifth

Table 8. Summary of Parameters and Questions Addressed

Parameter	Question Addressed
RMS of difference of maps	How well do annual mean maps agree?
RMS of seasonal cycle	What is the seasonal variability over the globe?
Δ_n	What is the map of differences for $\Phi_n(t)$?
$\|\|\Delta_n(x)\|\|$	What is RMS of difference maps?
$\|\|y(x, t) - Y(x, t)\|\|$	What is RMS of seasonal cycles?
H_{mn}	How well do PCs agree?
K_{mn}	How well do geographic patterns (EOFs) agree?

phase of the Coupled Model Intercomparison Project (CMIP5), whereby the various parameters could be evaluated for each model. This technique may be used to isolate quickly regions where clouds are biasing the seasonal cycle. This application of the method dealt with the seasonal cycles of a model and a measurement data set. The principal components of any two data sets could be compared in like manner. A major advantage of the method is that it provides a comparison over the entire domain, e.g., time or space, rather than a point comparison.

Appendix A

A1. Relations Among PCs

[47] The principal components describe the time behavior of the system. To provide a quantitative comparison of the two sets of PCs, the correlation matrix of **H** for the two sets of principal components was defined by equation (10). The properties of **H** are now considered. Each set of PCs is a complete orthonormal basis set, and one can be transformed into the other by a rotation matrix H_{mn}:

$$\varphi_m(t) = \sum_{n=1}^{12} \mathbf{H}_{mn}\Phi_n(t). \tag{A1}$$

By orthogonality,

$$\mathbf{H}_{mn} = \sum_{t=1}^{12} \varphi_m(t)\Phi_n(t)/12. \tag{A2}$$

Equation (A2) shows that the rotation matrix H_{mn} is the correlation matrix between $\varphi_n(t)$ and $\Phi_n(t)$ as defined by equation (10). Conversely,

$$\Phi_n(t) = \sum_{m=1}^{12} \mathbf{H}_{mn}\varphi_m(t). \tag{A3}$$

From equation (A2) and orthonormality of $\varphi_n(t)$ and $\Phi_n(t)$ it follows that

$$\sum_{k=1}^{12} \mathbf{H}_{mk}\mathbf{H}_{nk} = \delta_{mn}, \tag{A4}$$

that is, H_{mn} is an orthogonal transformation. Note that

$$\|\Phi_n(t) - \varphi_n(t)\|^2 = 2(1 - H_{nn}). \tag{A5}$$

where $\|\cdot\|$ is defined as the RMS in this paper and is a Euclidean measure in the vector spaces used here [*Naylor and Sell*, 1971]. The H_{mn} matrix is thus a measure of the level of agreement of the two data sets in time response. For perfect agreement of the two sets, H_{mn} is the identity matrix.

[48] Most of the information for evaluating the agreement of the data sets is in the diagonal terms of the H_{mn} matrix. Table 6 shows the diagonal terms of the H_{mn} matrix for ASR and OLR over land and over ocean. The H_{11} term equals 1 out to three decimal places, showing excellent agreement of the first principal components of the two data sets. As a consequence of equation (A4), if H_{mm} is nearly one, the remaining terms on row m are very small. As m increases, the off-diagonal terms of H_{mn} increase as the PCs become dominated by noise and are thus less similar between the two data sets. Table 2 shows that the RMS of the seasonal cycles of OLR over ocean is smaller than ASR by a factor of about 6. The smallness of the OLR variations makes the PCs more susceptible to errors, thus they do not agree as well as do the ASR PCs.

A2. Relations Among EOFs

[49] The empirical orthogonal components describe the spatial behavior of the system, analogous to the PCs description of the time behavior. One can similarly define a transformation matrix that relates the EOFs of the two data sets, enabling the quantification of the agreement of the spatial behavior of the two data sets in a manner similar to that for the PCs. The matrix K_{mn} may be defined by

$$\mathbf{K}_{mn} = \sum_x w_x \psi_m(x)\Psi_n(x) \tag{A6}$$

in analogy to equation (A2). The K_{mn} is the correlation matrix for the EOFs. The analogues to equations (A1), (A3), (A4) and (A5) also apply to the K_{mn}:

$$\psi_m(t) = \sum_{n=1}^{12} \mathbf{K}_{mn}\Psi_n(t), \tag{A7}$$

$$\Psi_n(t) = \sum_{m=1}^{12} \mathbf{K}_{mn}\psi_m(t), \tag{A8}$$

$$\sum_{k=1}^{12} \mathbf{K}_{mk}\mathbf{K}_{nk} = \delta_{mn}, \tag{A9}$$

and

$$\|\Psi_n(t) - \psi_n(t)\|^2 = 2(1 - \mathbf{K}_{nn}). \tag{A10}$$

The diagonal terms of K_{mn} provide a measure of how well the maps agree for each term. The first four diagonal terms of the matrices are listed in Table 7 for the cases of ASR and OLR over land and ocean. For ASR over both land and ocean, K_{11} is 0.996, which says the geographic distributions of ASR for the first PC, or the annual cycle of the energy absorbed into the system, agree practically perfectly. Likewise, for $n = 2$, the semiannual cycle, the agreement is good. For $n \geq 3$, K_{nn} has decreased so that the higher-order terms do not agree well, but these terms are small. For OLR over land and ocean the maps agree well for the first term, but over ocean K_{nn} drops quickly for $n > 1$. The geographical distribution of the annual cycle of the energy sink is described well, but the higher-order terms, which have little power, are not well described.

A3. Relations Between Δ_n Maps and EOFs

[50] The Δ_n are the maps of the differences of the two data sets projected onto the Φ_n, so the Δ_n are related to the maps of the projections of the two data sets onto their PCs, that is, the empirical orthogonal functions $\Psi_n(x)$ and $\psi_n(x)$. The behavior of the $\|\Delta_n\|$ can be explained in terms of the H_{mn} and K_{mn} matrices. Starting with equation (7) and using equations (4) and (6), it can be shown that

$$\|\Delta_n(x)\|^2 = \sum_m \left(r_m^2\,\mathbf{H}_{mn}^2 - 2r_m R_n \mathbf{H}_{mn}\mathbf{K}_{mn} + R_n^2\right). \tag{A11}$$

Δ_n will vanish if and only if (i) $r_n = R_n$, (ii) $H_{mn} = \delta_{mn}$ and (iii) $K_{mn} = \delta_{mn}$ for all m and n. For Δ_n to be zero for a

given n, the n-th magnitudes, time histories, and geographical distributions must be the same for both data sets. In equation (A11), for a large n, even though the r_n and R_n are small, the larger value of r_m for a small m can cause a significant difference if the H_{mn} is not small. For example, a large H_{31} will cause Φ_1 to project into Δ_3 with the RMS of term 1. Also, if K_{nn} is small for large r_n, indicating that the n-th maps do not agree, the second term will not reduce the right hand side of equation (A11), resulting in a large $\|\Delta_n(x)\|$.

[51] The model results can be expressed on the basis of the time response of EBAF by defining:

$$\omega_n(x) = \sum_{t=1}^{12} \Phi_n(t) y(t,x). \tag{A12}$$

The $\omega_n(x)$ are a set of maps of model results which are directly comparable to the EBAF EOFs. By use of equations (6) and (A1), the $\omega_n(x)$ can be expressed in terms of the $\psi_n(x)$ basis more generally:

$$\omega_n(x) = \sum_{m=1}^{12} \mathbf{H}_{mn}\, r_m\, \psi_m(x). \tag{A13}$$

The terms $H_{mn}r_m$ are the elements of the rotation matrix of the PCs weighted by the RMS of the model principal components. These $H_{mn}r_m$ coefficients project the model spatial distributions onto the EBAF time histories, that is, the $H_{mn}r_m$ tell how much each model map $\psi_n(x)$ mixes with the EBAF time variation $\Phi_n(t)$. Because H_{nn} is so close to 1 for the first four n's for ASR,

$$\omega_n(x) \cong r_n \cdot \psi_n(x) \text{ for } n = [1,\ 2,\ 3,\ 4]. \tag{A14}$$

Using equation (A13) to define $\omega_n(x)$, and equation (4) for $Y(x, t)$, $\Delta_n(x)$ may be expressed by using equation (7) as

$$\Delta_n(x) = \omega_n(x) - R_n \Psi_n(x). \tag{A15}$$

Therefore, for ASR over land, the $\Delta_1(x)$ map is approximated by

$$\Delta_1(x) \cong 75.2\psi_1(x) - 70.6\Psi_1(x). \tag{A16}$$

The annual cycles are described by the first PCs, which are practically the same for both data sets, and their EOFs, which are also very similar. The major difference between the ASR of the two data sets is that the annual cycle of the model has a greater magnitude than that of the EBAF. The semiannual cycles are represented by the second principal components and their EOFs and account for about 2.5% for each case for ASR over land. H_{mn} is nearly one for $n = m = 2$ and insignificant for higher n, so the contribution of $\psi_2(x)$ to $\omega_2(x)$ is an order of magnitude higher than any other term. The second term accounts for more than half of the variance of ASR that is not accounted for by the first term. Thus,

$$\Delta_2(x) \cong 12.5\psi_2(x) - 12.2\Psi_2(x). \tag{A17}$$

The remaining terms account for less than 2% of the variance of the data sets. There is little benefit in discussing terms of order higher than two, as they are very small and any changes to the model or to the data processing of CERES will make changes in the first two principal components which will overwhelm these small terms.

[52] **Acknowledgments.** We gratefully acknowledge support by the CERES program from the NASA Science Mission Directorate through Langley Research Center to Science Systems and Applications, Inc. The CERES EBAF data product was obtained from the NASA Langley Research Center CERES ordering tool at http://ceres.larc.nasa.gov/. We must also thank the Global Modeling and Assimilation Office of Goddard Space Flight Center for the GEOS-5 data set that was used in this paper. We are grateful to Kuan-Man Xu and the reviewer for their comments and suggestions. In particular, we have incorporated comments from a reviewer in the discussion of Figure 3. We thank the reviewer for this insight.

References

Barker, H. W., Z. Li, and J.-P. Blanchet (1994), Radiative characteristics of the Canadian Climate Centre second-generation general circulation model, *J. Clim.*, *7*, 1070–1091, doi:10.1175/1520-0442(1994)007<1070:RCOTCC>2.0.CO;2.

Barkstrom, B. R. (1990), Earth radiation budget measurements: Pre-ERBE, ERBE, and CERES, in *Long-Term Monitoring of the Earth's Radiation Budget*, *Proc. SPIE Int. Soc. Opt. Eng.*, *1299*, 52–60, doi:10.1117/12.21364.

Bony, S., H. Le Treut, J.-P. Duvel, and R. S. Kandel (1992), Satellite validation of GCM-simulated annual cycle of the Earth radiation budget and cloud forcing, *J. Geophys. Res.*, *97*(D16), 18,061–18,081, doi:10.1029/92JD01631.

Cess, R. D., et al. (1997), Comparison of the seasonal change in cloud-radiative forcing from atmospheric general circulation models and satellite observations, *J. Geophys. Res.*, *102*(D14), 16,593–16,603, doi:10.1029/97JD00927.

Chen, C.-T., and E. Roeckner (1996), Validation of the Earth radiation budget as simulated by the Max Planck Institute for Meteorology general circulation model ECHAM4 using satellite observations of the Earth Radiation Budget Experiment, *J. Geophys. Res.*, *101*(D2), 4269–4287, doi:10.1029/95JD03195.

Cullather, R. I., Harshvardhan, and K. A. Campana (1997), Climatology of cloud and radiation fields in a numerical weather prediction model, *Theor. Appl. Climatol.*, *57*, 11–33, doi:10.1007/BF00867974.

Harrison, E. F., P. Minnis, B. R. Barkstrom, V. Ramanathan, R. D. Cess, and G. G. Gibson (1990), Seasonal variation of cloud radiative forcing derived from the Earth Radiation Budget Experiment, *J. Geophys. Res.*, *95*(D11), 18,687–18,703, doi:10.1029/JD095iD11p18687.

Kiehl, J. T., J. J. Hack, and J. W. Hurrell (1998), The energy budget of the NCAR Community Climate Model CCM3, *J. Clim.*, *11*, 1151–1178, doi:10.1175/1520-0442(1998)011<1151:TEBOTN>2.0.CO;2.

Lin, W. Y., and M. H. Zhang (2004), Evaluation of clouds and their radiative effects simulated by the NCAR Community Atmospheric Model against satellite observations, *J. Clim.*, *17*, 3302–3318, doi:10.1175/1520-0442(2004)017<3302:EOCATR>2.0.CO;2.

Loeb, N. G., S. Kato, K. Loukachine, and N. M. Smith (2005), Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Terra satellite. Part I: Methodology, *J. Atmos. Oceanic Technol.*, *22*, 338–351, doi:10.1175/JTECH1712.1.

Loeb, N. G., S. Kato, K. Loukachine, N. M. Smith, and D. R. Doelling (2007), Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Terra satellite. Part II: Validation, *J. Atmos. Oceanic Technol.*, *24*, 564–584, doi:10.1175/JTECH1983.1.

Loeb, N. G., B. A. Wielicki, D. R. Doelling, G. L. Smith, D. F. Keyes, S. Kato, N. Manalo-Smith, and T. Wong (2009), Towards optimal closure of the Earth's top-of-atmosphere radiation budget, *J. Clim.*, *22*, 748–766, doi:10.1175/2008JCLI2637.1.

Minnis, P., et al. (2011), CERES Edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data, Part I: Algorithms, *IEEE Trans. Geosci. Remote Sens.*, *49*, 4374–4400, doi:10.1109/TGRS.2011.2144601.

Mlynczak, P. E., G. L. Smith, and D. R. Doelling (2011), The annual cycle of Earth radiation budget from Clouds and the Earth's Radiant Energy System (CERES) data, *J. Appl. Meteorol. Climatol.*, *50*, 2490–2503, doi:10.1175/JAMC-D-11-050.1.

Molod, A., L. Takacs, M. Suarez, J. Bacmeister, I.-S. Song, and A. Eichmann (2012), The GEOS-5 Atmospheric General Circulation Model: Mean climate and development from MERRA to Fortuna, *NASA Tech. Memo.*, *104606*, vol. 28, 115 pp.

Naylor, A. W., and G. R. Sell (1971), *Linear Operator Theory in Engineering and Science*, 624 pp., Holt, Rinehart and Winston, New York.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng (1982), Sampling errors in the estimation of empirical orthogonal functions, *Mon.*

Weather Rev., *110*, 699–706, doi:10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2.

Potter, G. L., and R. D. Cess (2004), Testing the impact of clouds on the radiation budgets of 19 atmospheric general circulation models, *J. Geophys. Res.*, *109*, D02106, doi:10.1029/2003JD004018.

Potter, G. L., D. C. Bader, M. Riches, A. Bamzai, and R. Joseph (2011), Celebrating two decades of the Program for Climate Model Diagnosis and Intercomparison, *Bull. Am. Meteorol. Soc.*, *92*, 629–631, doi:10.1175/2011BAMS3018.1.

Rienecker, M., et al. (2008), The GEOS-5 Data Assimilation System—Documentation of versions 5.0.1, 5.1.0, and 5.2.0, *NASA Tech. Memo.*, *104606*, vol. 27, 118 pp.

Smith, G. L., K. J. Priestley, N. G. Loeb, B. A. Wielicki, T. P. Charlock, P. Minnis, D. R. Doelling, and D. A. Rutan (2011), Clouds and Earth Radiant Energy System (CERES), a review: Past, present and future, *Adv. Space Res.*, *48*, 254–263, doi:10.1016/j.asr.2011.03.009.

Taylor, P. C., R. G. Ellingson, and M. Cai (2011), Seasonal variations of climate feedbacks in the NCAR CCSM3, *J. Clim.*, *24*, 3433–3444, doi:10.1175/2011JCLI3862.1.

Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee, G. L. Smith, and J. E. Cooper (1996), Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System experiment, *Bull. Am. Meteorol. Soc.*, *77*, 853–868, doi:10.1175/1520-0477(1996)077<0853:CATERE>2.0.CO;2.

Wilber, A. C., G. L. Smith, S. K. Gupta, and P. W. Stackhouse (2006), Annual cycles of surface shortwave radiative fluxes, *J. Clim.*, *19*, 535–547, doi:10.1175/JCLI3625.1.

Wilks, D. S. (1995), *Statistical Methods in the Atmospheric Sciences*, 467 pp., Academic, San Diego, Calif.

Yang, S.-K., Y.-T. Hou, A. J. Miller, and K. A. Campana (1999), Evaluation of the Earth radiation budget in NCEP–NCAR Reanalysis with ERBE, *J. Clim.*, *12*, 477–493, doi:10.1175/1520-0442(1999)012<0477:EOTERB>2.0.CO;2.